Quality policy and the role of assessment in work-integrated learning

JON YORKE[1]
Curtin University, Perth, Australia
LESLEY VIDOVICH
University of Western Australia, Perth, Australia

This paper examines higher education quality policy developments internationally (U.K., U.S.) and in Australia with respect to the role of learning standards and assessment in work-integrated learning. Whilst remaining located primarily within the Australian higher education context, the paper briefly identifies some of the more influential global drivers and identifies how they play a significant role in shaping national agendas. The second part of this paper traces the development of quality policy in Australian higher education with respect to learning standards, and in particular it focuses on the policy intention to develop 'direct measures' of learning, which, it is argued, is potentially detrimental to work-integrated learning (WIL). The third part of this paper analyses key tensions associated with learning standards in terms of the assessment practices used to warrant them, identifying specific challenges in the context of WIL. The final section of the paper draws together these issues, identifying a number of implications and opportunities for the assessment of WIL in the current Australian quality policy environment. *(Asia-Pacific Journal of Cooperative Education, Special Issue, 2014, 15(3), 225-239)*

Keywords: Quality, policy, learning standards, assessment, work-integrated learning

In higher education, assessment is more challenging than many would prefer to admit (Bloxham, Boyd, & Orr, 2011; Yorke, 2008), and these challenges are, if anything, even more profoundly experienced in the area of work-integrated learning (WIL). The field of WIL encompasses a diffuse range of on- and off-campus learning experiences, either in the workplace or in simulations of it, designed to enable students to integrate academic theory and workplace practice (Jackson, 2013). As Wilton (2012) pointed out, successful WIL experiences are seen by employers to offer key points of differentiation between graduates.

Judgments of student success arise from processes of assessment, and 'assessment' is taken here to mean the appraisal of student work in order to make a judgment of performance (Sadler, 2005). In the context of WIL, assessment might be conducted by those within the higher education institution or those external to it and this is an important point which will be picked up later. In either case, appraisals of student performance are increasingly made with respect to learning standards, which may be explicitly and/or implicitly defined. Learning standards are essentially key reference points that describe what students should know or can do (Bloxham et al., 2011; Price, 2005; Sadler, 2007).

Despite the significant attention directed towards assessment and learning standards by universities, it remains a challenging and contentious field. Students commonly identify assessment as an area in need of attention in their evaluations of taught programs (Jessop, El Hakim, & Gibbs, 2014; Scott, 2005; Williams & Kane, 2008). This sentiment is also reflected in

[1] Corresponding author: Jon Yorke, j.yorke@curtin.edu.au

institutional audits of quality. For example, Ewan (2009) in her thematic analysis of audits conducted by the Australian Universities Quality Agency noted that approximately a fifth of the audit recommendations for universities revolved around the need to improve or develop consistent assessment policies and practices within their institutional quality frameworks. More recently, echoing a broader point made by Yorke (2011), Natoli, Jackling, Kaider, and Clark (2013) described how the assessment practices in the discipline of accounting lagged behind their particular institution's new WIL policy. Some of the reasons for this were ascribed to problems of resourcing; others related to the way the policy authors and academic staff had different interpretations of some of the policy terminology (such as what was meant by a 'simulated environment'). Natoli et al. (2013) concluded by calling for university staff professional development and capacity building in the domain of WIL, coupled with more effective communication regarding the university's quality policy.

Conceptions of quality remain multiple, malleable, and highly contested. Harvey and Green (1993) offered a seminal definition of quality, suggesting that it could be described in any or all of the following terms: excellence; perfection; fitness for purpose; value for money; and transformation. Traditional conceptions of quality in higher education as 'excellent standards' (Vidovich, 2001) have much in common with Harvey and Green's first classification. Such standards may be maintained (and advanced) through processes of *quality improvement,* and proven through processes of *quality assurance.* Vidovich (2009) has suggested that the balance of power has, over time, shifted away from 'improve'(internally oriented critical reflections) and towards the 'prove' (externally oriented accountability), an observation also made elsewhere (Ranson, 2003; Reid, 2009). A heightened focus on accountability is evident in quality policy discourses on a global scale, and these, it is argued here, set an important context for the assessment of work-integrated learning.

GLOBALIZATION AND THE RISE OF QUALITY POLICIES

Higher education is increasingly seen as the engine driving economic productivity in a global knowledge society, and many countries have set ambitious participation targets for higher education (Bradley, Noonan, Nugent, & Scales, 2008). However, the transition from elite to mass participation in higher education has augmented the cost burden on nation states, raising questions of quality and triggering concerns over value for money. The Organisation for Economic Cooperation and Development (OECD), in its landmark report titled *Tertiary Education for the Knowledge Society* identified a number of global trends in quality assurance, recognizing the adoption of new public management approaches across a number of OECD governments with an emphasis on "leadership principles, incentives and competition" (Santiago, Tremblay, Basri, & Arnal, 2008, p. 260). Here, quality assurance was linked to economic growth through the signals sent to labor markets by high quality graduates who are able to participate and compete in a global marketplace. Contemporary quality policies seek to assure learning standards in a globally competitive context, and to provide a level of protection for higher education consumers (students, employers) from substandard providers.

The OECD initiated a major project titled *Assessment of Higher Education Learning Outcomes* (AHELO) in January 2010. Within this three year project, 248 institutions in 17 countries took part (OECD, 2012). This included the U.S. and Australia as full participants, with England taking part as an observer. AHELO sought to develop and apply standardized testing instruments to evaluate generic skills (critical thinking, analytical reasoning, problem solving and written communication) and some discipline-specific skills within the fields of economics and engineering. Contextual information regarding the background and learning environments was also collected as part of the project, which was designed to support processes of international benchmarking (OECD, 2011), in a similar way to the OECD's Program for International Student Assessment (PISA) benchmarking mechanism in the schooling sector. Significantly, AHELO focused on the 'academic' outcomes outlined above, leaving a number of aspects of WIL unexplored, and potentially marginalized. Even though AHELO analyses were confined to a limited subset of indicators in the generic skill strand, it was determined that "the instrument would require further consultation to provide evidence of content validity" (OECD, 2013a, p. 30) before it could be used as a basis for international comparisons. The final volume of the feasibility study report concluded that there was an ongoing need for AHELO type data, but the national costs were not known and "more data and analysis was still to be gained from the feasibility study" (OECD, 2013b, p. 45) before making any decision to proceed with a full scale implementation.

The growing focus on quality and standards, arguably forged by the OECD and other international organizations and networks, is evident in policy developments in countries across the globe, and in subsequent paragraphs these trends are discussed specifically in relation to the U.K., Europe, and U.S., before presenting the Australian policy context in the following section.

The U.K. has had a long standing interest in the quality and comparability of academic qualifications across and beyond the sector. Early work conducted by Johnes and Taylor (1990) reported a number of differences between institutional degree classifications in an attempt to 'benchmark' or compare standards across institutions. In 1999, 'subject benchmark statements' were brought to prominence following a landmark review of U.K. Higher Education titled *Higher Education in the Learning Society* (National Committee of Inquiry into Higher Education, 1997). Subject benchmark statements describe the threshold and typical sets of attributes, skills and capabilities that a graduate of a particular discipline would be expected to have. These subject benchmark statements, setting out broad expectations for graduates of the discipline, were developed in response to challenges about standards. It is pertinent to note that these were originally termed benchmark *standards*, however, by the time the first benchmarks were published in May 2000 they had been relabeled as benchmark *statements*, a move which prompted comment that this "change recognized the failure of the process to clearly define explicit standards for all subjects" (Rust, Price, & O'Donovan, 2003, p. 148). These statements are not intended to specify a detailed curriculum or favor particular assessment approaches. Instead, they are intended to inform and assist those involved in program design and review through an established

consensus on the nature of standards in that discipline (Quality Assurance Agency for Higher Education, 2010).

In Europe, the signing of the Bologna Declaration in 1999 by ministers from 29 countries signaled a commitment to work towards comparable degree specifications with broad outcomes defined at bachelor, masters and doctoral levels (European Higher Education Area, 1999). By 2000, the European 'Tuning' project was linking the Bologna declaration to activities in the education sector, ultimately proceeding to produce a 'tuning process' and a set of associated tools (Tuning Management Committee, 2006) to enhance inter-national alignment of degree standards. For example, the Competences in Education and Recognition Project produced *A Tuning Guide to Formulating Degree Qualification Profiles*, which included references to program learning outcomes, the purpose of which was to "describe accurately the verifiable learning achievements of a student at a given point in time" (Lokhoff et al., 2010, p. 22).

In the U.S., an increased focus on learning outcomes, their measurement and their comparability was also becoming more apparent at the turn of the millennium. In the school sector the enactment of *No Child Left Behind* in January 2002 had introduced standardized testing for all students on an annual basis, with severe measures for those schools that failed to demonstrate adequate yearly progress (Hursh, 2008). In the higher education area, the Spellings Commission published *A Test of Leadership: Charting the Future of U.S. Higher Education*, calling for increased accountability and suggesting that "higher education institutions should measure student learning using quality assessment from instruments such as, for example, the Collegiate Learning Assessment" (U.S. Department of Education, 2006, p. 24).

The Collegiate Learning Assessment had been released in 2002 by the Council for Aid to Education (a not-for-profit educational foundation) in order to measure critical thinking, analytical reasoning and written communication using a standardized test. However, this test does not include broader WIL related aspects such as teamwork, oral communication, civic engagement, ethical reasoning, or intercultural knowledge and competence. Partly in reaction to what were perceived as 'narrow' quality indicators and partly in anticipation of increased accountability foreshadowed by the Spellings Commission, the Association of American Colleges and Universities embarked on a three year mission in 2007 to develop *Valid Assessments of Learning in Undergraduate Education*. A key outcome of this project was to collaboratively develop and agree a set of detailed standards for 15 essential learning outcomes, including those associated with the assessment of WIL, expressed at the level of the graduate (Association of American Colleges and Universities, 2006). More recently, the American Lumina Foundation launched the *Degree Qualifications Profile* (DQP), which "illustrates clearly what students should be expected to know and be able to do once they earn their degrees" (Lumina Foundation, 2011, p. 1). The DQP describes five areas of learning, which are "Broad, Integrative Knowledge; Specialized Knowledge; Intellectual Skills; Applied Learning, and Civic Learning" (Lumina Foundation, 2011, p. 4).

In all, over a number of years, U.K., European and American higher education quality policies and various stakeholders have given increasing attention to learning standards and, in particular, their comparability across disciplines, institutions, and regions. More recently this has extended to include international comparisons through the work of the OECD (2012, 2013a, 2013b). In many settings there is an interest in furthering WIL (the American DQP is a good example of this). However, when it comes to the quality assurance of learning, standards tend to be more narrowly defined and these often neglect aspects of WIL. This is a particularly ironic state of affairs given that quality policies aspire to improve employability and advance economic productivity. These evolving policy developments are also readily recognizable in the Australian context, as outlined in the next section.

QUALITY POLICY DEVELOPMENTS IN AUSTRALIA

In Australia, the focus on 'quality' in higher education sharpened in the early 1990s with Australia's first official quality policy (Baldwin, 1991). This policy largely constructed quality in terms of 'excellent standards' in universities, focusing on institutional processes. This view was to gradually transition towards discourses of quality assurance (Vidovich, 2001). In 1999, Minister Kemp pointed to the inability to compare Australian standards with other countries (Kemp, 1999), foreshadowing heightened interest in international competitiveness in higher education in the first decade of the new millennium. By early 2008, following the election of an Australian Labor Government, the Minister for Education (Gillard) initiated a major review of the sector, under the aegis of a panel chaired by former Vice-Chancellor Denise Bradley with a remit to examine whether higher education was "meeting the needs of the Australian community and economy" (Bradley et al., 2008, p. 205) in the international marketplace. The final report of the panel, titled *Review of Australian Higher Education* and known as the 'Bradley report', was released in December 2008, heralding significant change and a renewed emphasis on quality outcomes from the higher education sector. It highlighted the need for Australia to invest to increase the proportion of the population with a bachelor degree, and argued that stronger accountability was needed to improve and assure the quality of those graduates. Underpinning these conclusions lay a belief that Australia was losing ground against other countries, thereby being positioned at a competitive disadvantage globally. The Bradley report ultimately made 46 recommendations to the Federal Government (Bradley et al., 2008), and in the context of WIL, Recommendation 23 is particularly relevant. Recommendation 23 firmly positioned learning standards within the umbrella of the new quality assurance arrangements, arguing that "a set of indicators and instruments to directly assess and compare learning outcomes" was needed, together with "a set of formal statements of academic standards by discipline along with processes for applying those standards" (Bradley et al., 2008, p. 137).

The Australian Government accepted the majority of Bradley's recommendations, and in 2009 a landmark policy framework was put forward. Positioned as an overarching ten year vision, *Transforming Australia's Higher Education System* articulated a comprehensive response

to the Bradley review, ushering in what was labelled "A New Era of Quality in Australian Tertiary Education" (Australian Government, 2009, p. 31).

In the four years that have followed the release of *Transforming Australia's Higher Education System*, various Government policy consultations have sought to establish the direct measures of learning proposed by Bradley et al. (2008), but these moves were vigorously resisted by the sector. Proposals to reward universities for the quality of learning outcomes via performance funding were challenged, and eventually dropped. Various standardized tests of learning standards were proposed as performance indicators, such as the Graduate Skills Assessment and the Collegiate Learning Assessment (from the U.S.), but these suggestions were met with considerable challenge by a sector that saw these indicators as flawed, narrowly defined, and irrelevant to the Australian context. Proposals for performance indicators of learning standards were finally dropped following another round of consultation (Advancing Quality in Higher Education Reference Group, 2012). In their place, and to assure that the generic skills of graduates were able to meet the needs of employers, the reference group proposed a literature review and scoping study to investigate the feasibility of an employer satisfaction survey.

During this period, a number of projects relating to learning standards emerged. The former Australian Learning and Teaching Council (ALTC) initiated the Learning and Teaching Academic Standards project, which drew together disciplinary groupings, led by 'Discipline Scholars', who were experienced academics at a professorial level (ALTC, 2010). The key deliverable for Discipline Scholars was the production of a document containing 'Threshold Learning Outcomes', as a step towards articulating the minimum standards for graduation in that discipline. Other projects emerged with a view to verifying standards through processes of external comparison, such as the Quality Verification System (Group of Eight, 2011).

In 2013, a new consultation was released by the Higher Education Standards Panel, an expert advisory body established to provide independent advice to the Minister(s) responsible for tertiary education. This consultation sought comment from the sector on a new proposal that abandoned quantitative performance measures of learning standards in favor of judgment-based processes combined with periodical peer review (Australian Government, 2013). This approach would represent a policy turn that could potentially create an environment less hostile to the assessment of WIL. However, at the time of writing (December 2013), this policy discussion was ongoing and no decision had been reached.

Taken together, these international and Australian quality policy developments set an important context for the assessment of WIL. The discussion so far indicates that quality is becoming increasingly defined with an eye to international competitiveness, based on performance indicators that are often narrowly defined. There has been significant contestation regarding the assurance of learning standards within the sector more broadly, but policy deliberations to date have only marginally addressed issues specific to the assessment of WIL. Governments in the U.K., U.S. and Australia have struggled to establish 'measures' of learning in formal higher education institutions where the terrain is

comparatively well defined and controlled. In the diverse and less centrally controlled domain of WIL the issue of quality and standards becomes even more difficult. The following section explores some of the reasons why quality assurance of assessment in the WIL context is particularly challenging.

LEARNING STANDARDS, QUALITY AND WIL

The assessment of student performances in the light of learning standards is increasingly seen as central to institutional quality, but the problems associated with standards-based grading decisions are long-standing (Bloxham et al., 2011; Newstead & Dennis, 1994; Price, Carroll, O'Donovan, & Rust, 2011; Woolf, 2004; Yorke, 2008). Bloxham et al. (2011) suggested that there are two oppositional theoretical frameworks for assessment at the heart of arguments regarding learning standards. On the one hand, positivist views of assessment suggest that standards can be objectively defined with precision, and that achievement can be reliably *measured* against those standards. On the other hand, interpretative views see standards as those broader, tacit, normative, and consensually established *judgments*, or 'rules of the discipline'. There is significant contestation between these opposing positions with proponents of detailed exposition of learning standards arguing for approaches more consistent with specification and measurement, whilst others at the opposite end of this continuum call for holism and judgment. The following sections briefly address three key challenges for assessment in a WIL context. These challenges relate to the difficulty of defining and specifying learning standards; the problems with applying those standards to consistently appraise work through grading practices; and the differences in the way results from WIL and non-WIL contexts are treated. Each is examined in turn.

DEFINING AND SPECIFYING LEARNING STANDARDS

In the OECD's *Tertiary Education for the Knowledge Society*, Santiago et al. (2008, p. 312) took up the link between learning outcomes and quality, asserting that "in the absence of objective measures of learning outcomes, there is no way for students to judge the quality of TEIs [tertiary education institutions] except by reputation". However, these objective measures of learning standards may be somewhat elusive, and their pursuit may lead to undesirable consequences (Knight, 2002a; Sadler, 2007). These challenges are, if anything, more acutely experienced in the field of WIL (Yorke, 2011).

Clearly defined learning standards, coupled with 'transparent' assessment processes, are intended to improve consistency, thereby reducing the arbitrariness of staff decisions and rendering matters open to scrutiny by other interested parties (Boud & Associates, 2010). Furthermore, Stowell (2004) suggested that clear specification is needed in order to address issues of student equity. As Sadler (2009) has argued, the provision of clear criteria has become established 'best practice', to the point where detailed expositions of standards are considered to be mandatory in some universities.

However, the over-zealous specification of learning standards has been subject to sharp and persistent criticism by a number of scholars such as Sadler (2007, 2008, 2009) who pointed to the various ways in which students are 'short changed' by tightly defined standards. Despite having well-documented outcomes and criteria, "final-year students were not clear about goals and standards" in a recent study of 23 degree programs in eight U.K. universities (Jessop et al., 2014, p. 82). In fact, standards that are too closely prescribed may even be counter-productive to learning. Discussing assessment matters in a workplace context, Torrance (2007) reported that the explicit specification of learning standards in the interests of clarity and transparency had demonstrably led to narrow and instrumental responses on the part of learners. This position is especially unfortunate given that higher education aspires to allow for (and encourage) a variety of equally acceptable approaches to a particular task, and this diversity is particularly relevant in the case of WIL.

Knight (2002b, p. 280) in his analysis of assessment practices drew together Eisner's (1985) conceptions of 'connoisseurship' with learning theory and psychometrics to assert that in higher education, "benchmarks, specifications, criteria and learning outcomes do not and cannot make summative assessment reliable, may limit its validity and certainly compound its costs". Connoisseurship describes the subjective but skillful judgments made by experienced practitioners who have progressively immersed themselves in the discipline over time as they progress in a community of academic practice (Lave & Wenger, 1991). A similar line is taken by Jawitz (2009), who drew on Bourdieu's conceptions of 'habitus' to suggest that staff gradually become more familiar with the tacit criteria and standards through social processes of consensus building over time. However, as Sadler (2009) argued, holistic judgments based on connoisseurship may be at odds with the grade or mark produced by following a closely prescribed set of standards for each of the criteria. On this point, Trede and Smith (2014, p. 165) observed that when workplace assessors' judgment-based mark did not align with the mark that was obtained from the use of a standard assessment form, they "would not act on their judgment but reverted back to dominant practices as defined by the material objects they had been given". In other words, the skillful judgments of connoisseurship were overruled by the specifications contained in the assessment forms. Whilst these issues raise questions for assessment practices more broadly, they are far more problematic for WIL, especially given that assessors external to the higher education institution are likely to be peripheral to many institutional processes.

GRADING PRACTICES

In seminal work conducted by Newstead and Dennis (1994), the reliability of grades awarded by highly experienced academics in the field of psychology was shown to be poor, even where the task domain and expected learning standards were clearly specified. Since then, this has been repeatedly observed in successive studies (Bloxham, 2009; Newstead, 2002; Orr, 2007; Price et al., 2011). Knight (2006) referred to the local practices of assessment to describe how grading decisions are made in a particular context, arguing that they reflect both the nature of the assessment task and the circumstances in which the assessor made

their judgment. It may be extremely difficult to achieve consistency on grades/marks awarded for individual assessment tasks, even where learning standards are comprehensively and exhaustively defined. Student performance in a workplace setting, which can be a comparatively 'uncontrolled' assessment environment, is impacted by a diverse set of variables (Cooper, Orrell, & Bowden, 2010). From a quality assurance point of view, these problems become almost intractable in the 'messier' world of WIL.

The consistency of grading judgments may be improved through processes of moderation, including post-hoc analyses of results or more holistic approaches that include *a priori* activities such as consensus building pre-marking meetings (Oliver, Lawson, & Yorke, 2008). Post-hoc moderation approaches to the quality assurance of learning standards include 'double marking' where a second assessor appraises the work, and a variety of strategies may be employed such as random sampling, purposive sampling according to pre-set rules, or some combination of both. Second marking may also be conducted 'blind' where first marker comments are made unavailable to successive markers. However, Bloxham (2009) has warned of the 'false promise' of moderation for the purposes of quality assurance, pointing to power dynamics in marking and the imperfect way in which disagreements between markers are rationalized. Often, she argued, the different marks given by two disagreeing markers are simply averaged, an observation borne out in empirical studies conducted by Orr (2007). The averaging of differing marks is perhaps a disservice to the judgments of both markers.

Furthermore, many approaches to moderation assume the presence of an assessment 'artefact' of some sort, such as an essay, reflective report, or a video of practice. In the WIL context, such artefacts are not necessarily always readily available, and the level of intervention needed to acquire such evidence may serve to detract from the learning experience. There are a number of difficulties associated with the moderation of mentors' judgments of practice based on a series of observations in a variety of settings over time, for example. For these and other reasons, it is perhaps unsurprising that institutions elect to retain responsibility for WIL related assessment decisions to the extent they do (Ferns & Moore, 2012).

GRADING IN WIL AND NON-WIL CONTEXTS

The discussion so far has highlighted a number of problems with learning standards, from their definition to their application. Higher education outcomes have historically been determined from an aggregation of individual assessment task grades or marks. Assessment results can be described in a number of ways, and there are strengths and weaknesses in each approach (Yorke, 2008). However, it is not uncommon for the assessment of WIL to be treated differently to assessment in other (non-WIL) areas within the university.

In the Australian university setting, individual assessment task results are commonly expressed as a percentage mark. However, the awarding of numerical percentages in the WIL setting is less common (Reddan, 2013) where results are more commonly cast as a

binary pass/fail, or located in stratified bands of performance using single letter grades (such as A, B, C) or descriptive terms (such as 'pass' or 'distinction'). The mathematical summation of different task results also poses a problem for the institution and there are widely reported statistical weaknesses encountered when individual assessment results are aggregated to generate an overall result (Knight, 2002a; Woolf, 2004; Yorke, 2008).

These issues pose a number of problems for the assessment of WIL, in that differing grading practices may instantiate different 'rules of engagement' with students. For example, Reddan (2013) reported that some students felt they would tend to put less effort into ungraded WIL activities. Furthermore, where WIL tasks are graded, the weighting of assessment accorded to WIL may be marginalized. Natoli et al. (2013, p. 80) remarked on the low apportioning of marks given to WIL activities, which signaled their lower status in comparison to other activities, pointing out that this risked downplaying the "depth and significance of any WIL learning [sic]". These are endemic problems of assessment but they become heightened in a WIL context.

IMPLICATIONS FOR THE ASSESSMENT OF WIL

As we have outlined, the assessment of tasks in the workplace, or close simulations of it, is somewhat 'messy'. The assessment of WIL often contains unpredictable aspects, and students often have to complete their task with variable, incomplete or inaccurate data. The 'scaffolding' of WIL experiences can be qualitatively different to that in non-WIL settings (see, for example, the discussion by Hodges, Eames, & Coll, 2014). Learning standards in relation to WIL are not easy to define, and they are decidedly difficult to apply consistently given the inherent variation within the domain. These issues represent a conundrum for Australian higher education institutions given the current Government policy focus on the development of direct measures of learning, assayed through indicators that have been up until now fairly narrowly defined. However, this problem is not new, nor is it unique to Australia. This final section summarizes four key issues that seem to be central to the assessment of WIL, and briefly sketches out some possibilities for future work.

Firstly, to an extent, assessment practices in WIL have lagged behind developments in the provision of WIL, and there remains an ongoing need for professional development opportunities for those assessing in a WIL context. This appears to be just as important for those engaged with WIL within the institution as well as those external to it. External participants in WIL programs often lack time and resources to engage with assessment issues, and this may be more acute for small to medium enterprises.

Second, the aggregation of student results from WIL and non-WIL assessment tasks should be organized to avoid inadvertently downplaying the status of WIL activities. Institutional quality and assessment policies need to be sensitive to this risk, and closer articulation of these activities may prove fruitful. For example, institutions often develop a separate policy to help emphasize new initiatives (see, e.g., Natoli et al., 2013) but more may be gained by embedding WIL in existing policies. Some universities have moved, at least in part, away

from pass/fail binary judgments of WIL towards the use of grades and found the outcome to be positive (Reddan, 2013). However, a parity of esteem will not be adequately established until both WIL and non-WIL judgments carry equal status. Given the limitations of percentage marking in a number of disciplines there is an emerging argument for harmonizing on a common grading format across WIL and non-WIL contexts, although this would need to be nuanced to the institutional and disciplinary setting.

Third, whilst it is difficult to define and apply learning standards with fine precision in many areas of higher education, this does not imply that a standards-based approach is inappropriate. Broadly consistent judgments of learning standards (and, by extension, quality) are possible, but both explicit and tacit knowledge is required if standards are to be communicated effectively in any direction within or between students or staff. With respect to the assurance of quality, assessor 'calibration' exercises (Sadler, 2012) offer a means of demonstrating consistency in the way learning standards are used without recourse to time consuming moderation activities.

Finally, there are significant risks for WIL if quality (at an institutional or national level) is construed using narrowly defined performance indicators that do not encompass those dimensions that are critical for success, but difficult to measure. The use of narrowly defined conceptions of learning may serve to marginalize other aspects such as ethical practice or working with others if such aspects fly under the radar of quality performance indicators. There may be far reaching implications for equity and diversity in higher education if quality and accountability is to be conflated with standardized testing (De Lissovoy & Mclaren, 2003).

Sharply defined performance indicators that are subject to public comparison may reorientate activities towards supporting those which have impact on these measures (Ball, 2012), a situation which may have damaging consequences. Appearances of poor performance (based on potentially flawed quality indicators) can have far reaching implications for institutions. One stark example of this is provided by Salmi (2009), who reported that demand for courses evaluated positively by the Brazilian Provão (a national mandatory assessment of graduate outcomes in certain fields of study) rose by some 20% whilst the demand for courses with a negative Provão assessment reduced by 41%; a high stakes test, indeed.

In short, a considerable body of literature suggests that direct, comparable and reliable measures of learning are beyond our current reach. If so, this suggests that quality policies ought to focus on improving judgments, shifting the focus more towards the establishment of a "controlled reputational range" (as cited in Brown, 2010, p. 3), an approach that is more tolerant of diversity.

CONCLUDING DISCUSSION

Massaro (2010, p. 23) suggested that quality policies are increasingly concerned with ensuring that education qualifications are globally transportable, "comparable in standard from one country to another", and quality assured through "reports to society that are comprehensible to it". These potentially competitive aspirations are clearly evident in the international and Australian policy agendas to advance economic prosperity through higher education outcomes. However, in the Australian context it is currently unclear whether the Government will continue to seek direct 'measures' of learning in future iterations of quality policy, or whether the softer accountabilities recently proposed by the Higher Education Standards Panel will prevail.

This paper adds to the body of literature arguing that quality should not be exclusively cast in terms of a narrowly defined set of indicators (at either national or institutional level). Alongside other weaknesses, narrow definitions of quality could be particularly detrimental to WIL. There is arguably more to be gained through the development of judgments, supported and appraised by processes of peer review, and underpinned by a philosophical acceptance that assessment in higher education is somewhat imprecise.

There are difficult challenges for the assessment of WIL embedded in the issues outlined here. In the Australian quality policy context, the discussion may be turning from measurement towards judgment, and the removal of narrow performance indicators would provide a policy environment that could potentially be sympathetic to the assessment of WIL. Admittedly, progress is likely to be neither easy nor swift. However, if it is agreed that the practices of WIL form a crucially important part of students' higher education, then much more attention needs to be given to assessment in this domain if we are to support and realize national policy intentions to develop highly employable graduates.

REFERENCES

Advancing Quality in Higher Education Reference Group. (2012). *Development of performance measures*. Canberra: Department of Innovation, Industry, Science and Research and Tertiary Education. Retrieved from http://www.innovation.gov.au/highereducation/Policy/Documents/AQHEreport.pdf

Association of American Colleges and Universities. (2006). *Statement on Spellings Commission Report*. Retrieved from http://www.aacu.org/about/statements/Spellings9_26_06.cfm

Australian Government. (2009). *Transforming Australia's higher education system*. Canberra, Austarlia: Department of Education, Employment, Workplace Relations and Social Inclusion.

Australian Government. (2013). *Draft standards for course design and learning outcomes*. Canberra, Australia.

Australian Learning and Teaching Council. (2010). Learning and teaching academic standards project: Final report. Strawberry Hill, New South Wales, Australia.

Baldwin, P. (1991). *Higher education: Quality and diversity in the 1990's*. Canberra: Australian Government Publishing Service.

Ball, S. J. (2012). Performativity, commodification and commitment: An I-spy guide to the neoliberal university. *British Journal of Educational Studies, 60*(1), 17-28. doi: 10.1080/00071005.2011.650940

Bloxham, S. (2009). Marking and moderation in the UK: False assumptions and wasted resources. *Assessment & Evaluation in Higher Education, 34*(2), 209-220. doi: 10.1080/02602930801955978

Bloxham, S., Boyd, P., & Orr, S. (2011). Mark my words: The role of assessment criteria in UK higher education grading practices. *Studies in Higher Education, 36*(6), 655-616.

Boud, D., & Associates. (2010). Assessment 2020: Seven propositions for assessment reform in higher education. Retrieved from http://www.olt.gov.au/system/files/resources/Assessment%202020_final.pdf

Bradley, D., Noonan, P., Nugent, H., & Scales, B. (2008). *Review of Australian higher education: Final report*. Canberra: Department of Education, Employment and Workplace Relations.

Brown, R. (2010). Comparability of degree standards? Retrieved from http://www.hepi.ac.uk/455-1838/Comparability-of-degree-standards.html

Cooper, L., Orrell, J., & Bowden, M. (2010). *Work-integrated learning: A guide to effective practice.* USA: Routledge.

De Lissovoy, N., & Mclaren, P. (2003). Educational 'accountability' and the violence of capital: A Marxian reading. *Journal of Education Policy, 18*(2), 131-143.

Eisner, E. W. (1985). *The art of educational evaluation: A personal view.* London: Falmer Press.

European Higher Education Area. (1999). *The Bologna declaration of 19 June 1999: Joint declaration of the European Ministers of Education*. Retrieved from http://www.ehea.info/Uploads/Declarations/BOLOGNA_DECLARATION1.pdf

Ewan, C. (2009). Learning and teaching in Australian universities: A thematic analysis of cycle 1 AUQA audits. Retrieved from http://www.olt.gov.au/system/files/resources/AUQA%20report.pdf

Ferns, S., & Moore, K. (2012). Assessing student outcomes in fieldwork placements: An overview of current practice. *Asia-Pacific Journal of Cooperative Education, 13*(4), 207-224.

Group of Eight. (2011). Go8 update. Retrieved from http://www.go8.edu.au/__documents/newsletters/2011/go8_newsletter_july2011.pdf

Harvey, L., & Green, D. (1993). Defining quality. *Assessment and Evaluation in Higher Education, 18*(1), 9-34.

Hodges, D., Eames, C., & Coll, R. (2014). Theoretical perspectives on assessment in cooperative education placements [Special issue]. *Asia-Pacific Journal of Cooperative Education, 15*(3), 189-205

Hursh, D. (2008). *High-stakes testing and the decline of teaching and learning: The real crisis in education.* Lanham, MD: Rowman & Littlefield.

Jackson, D. (2013). The contribution of work-integrated learning to undergraduate employability skill outcomes. *Asia-Pacific Journal of Cooperative Education, 14*(2), 99-115.

Jawitz, J. (2009). Learning in the academic workplace: The harmonization of the collective and the individual habitus. *Studies in Higher Education, 34*(6), 601-614. doi: 10.1080/03075070802556149

Jessop, T., El Hakim, Y., & Gibbs, G. (2014). The whole is greater than the sum of its parts: A large-scale study of students' learning in response to different programme assessment patterns. *Assessment & Evaluation in Higher Education, 39*(1), 73-88. doi: 10.1080/02602938.2013.792108

Johnes, J., & Taylor, L. (1990). *Performance indicators in higher education*. Buckingham: SRHE and Open University Press.

Kemp, D. (1999). *Quality assured: A new Australian quality assurance framework for university education*. (Workshop on the new quality assurance framework, Canberra, 10 December). Canberra.

Knight, P. (2002a). The Achilles' heel of quality: The assessment of student learning. *Quality in Higher Education, 8*(1), 107-115.

Knight, P. (2002b). Summative assessment in higher education: Practices in disarray. *Studies in Higher Education, 27*(3), 275-286.

Knight, P. (2006). The local practices of assessment. *Assessment & Evaluation in Higher Education, 31*(4), 435-452. doi: 10.1080/02602930600679126

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation.* Cambridge: Cambridge University Press.

Lokhoff, J., Wegewijs, B., Durkin, K., Wagenaar, R., González, J., Isaacs, A. K., . . . Gobbi, M. (2010). A Tuning guide to formulating degree programme profiles. Retrieved from http://core-project.eu/documents/Tuning%20G%20Formulating%20Degree%20PR4.pdf

Lumina Foundation. (2011). The Degree Qualifications Profile. Retrieved from http://www.luminafoundation.org/publications/The_Degree_Qualifications_Profile.pdf

Massaro, V. (2010). Cui bono? The relevance and impact of quality assurance. *Journal of Higher Education Policy and Management, 32*(1), 17-26. doi: 10.1080/13600800903440527

National Committee of Inquiry into Higher Education. (1997). *Higher education in the learning society*. Norwich: HMSO.

Natoli, R., Jackling, B., Kaider, F., & Clark, C. (2013). Mapping WIL activities in the curriculum to develop graduate capabilities: A case study in accounting. *Asia-Pacific Journal of Cooperative Education, 14*(2), 75-88.

Newstead, S. (2002). Examining the examiners: Why are we so bad at assessing students? *Psychology Learning & Teaching, 2*(2), 70-75.

Newstead, S., & Dennis, I. (1994). Examiners examined: The reliability of exam marking in psychology. *The Psychologist, 7*, 216-219.

Oliver, B., Lawson, K. J., & Yorke, J. D. (2008). *Moderation assessed: Policies and practices*. Paper presented at the ATN Assessment Conference 2008, Adelaide.

Organisation for Economic Cooperation and Development. (2011). Testing student and university performance globally: OECD's AHELO. from www.oecd.org/edu/ahelo

Organisation for Economic Cooperation and Development. (2012). AHELO feasibility study report: Design and implementation. *1*. Retrieved from http://www.oecd.org/edu/skills-beyond-school/AHELOFSReportVolume3.pdf

Organisation for Economic Cooperation and Development. (2013a). AHELO feasibility study report: Data analysis and national experiences. *2*. Retrieved from http://www.oecd.org/edu/skills-beyond-school/AHELOFSReportVolume2.pdf

Organisation for Economic Cooperation and Development. (2013b). AHELO feasibility study report: further insights. *3*. Retrieved from http://www.oecd.org/edu/skills-beyond-school/AHELOFSReportVolume3.pdf

Orr, S. (2007). Assessment moderation: Constructing the marks and constructing the students. *Assessment & Evaluation in Higher Education, 32*(6), 645-656. doi: 10.1080/02602930601117068

Price, M. (2005). Assessment standards: The role of communities of practice and the scholarship of assessment. *Assessment & Evaluation in Higher Education, 30*(3), 215-230. doi: 10.1080/02602930500063793

Price, M., Carroll, J., O'Donovan, B., & Rust, C. (2011). If I was going there I wouldn't start from here: A critical commentary on current assessment practice. *Assessment and Evaluation in Higher Education, 36*(4), 479-492.

Quality Assurance Agency for Higher Education. (2010). *Recognition scheme for subject benchmark statements*. Retrieved from http://www.qaa.ac.uk/Publications/InformationAndGuidance/Documents/recognitionscheme10.pdf

Ranson, S. (2003). Public accountability in the age of neo-liberal governance. *Journal of Education Policy, 18*(5), 459-480.

Reddan, G. (2013). To grade or not to grade: Student perceptions of the effects of grading a course in work-integrated learning. *Asia-Pacific Journal of Cooperative Education, 14*(4), 223-232.

Reid, I. (2009). The contradictory managerialism of university quality assurance. *Journal of Education Policy, 24*(5), 575-593.

Rust, C., Price, M., & O'Donovan, B. (2003). Improving students' learning by developing their understanding of assessment criteria and processes. *Assessment & Evaluation in Higher Education, 28*(2), 147-164. doi: 10.1080/02602930301671

Sadler, D. R. (2005). Interpretations of criteria-based assessment and grading in higher education. *Assessment and Evaluation in Higher Education, 30*(2), 175-194. doi: 10.1080/0260293042000264262

Sadler, D. R. (2007). Perils in the meticulous specification of goals and assessment criteria. *Assessment in Education: Principles, Policy & Practice, 14*(3), 387-392. doi: 10.1080/09695940701592097

Sadler, D. R. (2008). *Are we short-changing our students? The use of preset criteria in assessment.* Edinburgh: Centre for Teaching, Learning and Assessment.

Sadler, D. R. (2009). Indeterminacy in the use of preset criteria for assessment and grading. *Assessment and Evaluation in Higher Education, 34*(2), 159-179.

Sadler, D. R. (2012). Assuring academic achievement standards: From moderation to calibration. *Assessment in Education: Principles, Policy & Practice, 20*(1), 5-19. doi: 10.1080/0969594X.2012.714742

Salmi, J. (2009). The growing accountability agenda: Progress or mixed blessing? *Higher Education Management and Policy, 21*(1), 101-122.

Santiago, P., Tremblay, K., Basri, E., & Arnal, E. (2008). *Tertiary education for the knowledge society: Volume 1 : Special features: Governance, funding, quality*. Paris: OECD Publishing.

Scott, G. (2005). Accessing the student voice: Using CEQuery to identify what retains students and promotes engagement in productive learning in Australian higher education. Retrieved from http://pandora.nla.gov.au/tep/68139

Stowell, M. (2004). Equity, justice and standards: Assessment decision making in higher education. *Assessment & Evaluation in Higher Education, 29*(4), 495-510. doi: 10.1080/02602930310001689055

Torrance, H. (2007). Assessment as learning? How the use of explicit learning objectives, assessment criteria and feedback in post-secondary education and training can come to dominate learning. *Assessment in Education, 14*(3), 281-294.

Trede, F., & Smith, M. (2014). Workplace educators' interpretations of their assessment practices: A view through a critical practice lens. *Assessment & Evaluation in Higher Education, 39*(2), 154-167. doi: 10.1080/02602938.2013.803026

Tuning Management Committee. (2006). Tuning general brochure. Retrieved from http://www.unideusto.org/tuningeu/images/stories/documents/General_Brochure_final_version.pdf

U.S. Department of Education. (2006). *A test of leadership: Charting the future of U.S. higher education*. Washington, D.C.

Vidovich, L. (2001). That chameleon 'quality': The multiple and contradictory discourses of 'quality' policy in Australian higher education. *Discourse, 22*(2), 249-261.

Vidovich, L. (2009). You don't fatten the pig by weighting it: Contradictory tensions in the 'policy pandemic' of accountability infecting education. In M. Simons, M. Olssen & M. A. Peters (Eds.), *Re-reading education policies: studying the Policy Agenda of the 21st Century.* Rotterdam: Sense Publishers.

Williams, J., & Kane, D. (2008). Exploring the National Student Survey: Assessment and feedback issues. Higher Education Academy. Retrieved from http://www.heacademy.ac.uk/assets/documents/nss/nss_assessment_and_feedback_issues.pdf

Wilton, N. (2012). The impact of work placements on skill development and career outcomes for business and management graduates. *Studies in Higher Education, 37*(5), 603-620.

Woolf, H. (2004). Assessment criteria: Reflections on current practices. *Assessment & Evaluation in Higher Education, 29*(4), 479-493. doi: 10.1080/0260293031000168904 6

Yorke, D. M. (2008). *Grading student achievement in higher education: Signals and shortcomings*. New York: Routledge.

Yorke, M. (2011). Work-engaged learning: Towards a paradigm shift in assessment. *Quality in Higher Education, 17*(1), 117-130. doi: 10.1080/13538322.2011.554316

Asia-Pacific Journal of Cooperative Education
ISSN 1175-2882
www.apjce.org

About the Journal

The Asia-Pacific Journal of Cooperative Education publishes peer-reviewed original research, topical issues, and best practice articles from throughout the world dealing with Cooperative Education (Co-op) and Work Integrated Learning/Education (WIL).

In this Journal, Co-op/WIL is defined as an educational approach that uses relevant work-based projects that form an integrated and assessed part of an academic program of study (e.g., work placements, internships, practicum). These programs should have clear linkages with, or add to, the knowledge and skill base of the academic program. These programs can be described by a variety of names, such as work-based learning, workplace learning, professional training, industry-based learning, engaged industry learning, career and technical education, internships, experiential education, experiential learning, vocational education and training, fieldwork education, and service learning.

The Journal's main aim is to allow specialists working in these areas to disseminate their findings and share their knowledge for the benefit of institutions, co-op/WIL practitioners, and researchers. The Journal desires to encourage quality research and explorative critical discussion that will lead to the advancement of effective practices, development of further understanding of co-op/WIL, and promote further research.

Submitting Manuscripts

Before submitting a manuscript, please unsure that the 'instructions for authors' has been followed (www.apjce.org/instructions-for-authors). All manuscripts are to be submitted for blind review directly to the Editor-in-Chief (editor@apjce.org) by way of email attachment. All submissions of manuscripts must be in MS Word format, with manuscript word counts between 3,000 and 5,000 words (excluding references).

All manuscripts, if deemed relevant to the Journal's audience, will be double blind reviewed by two reviewers or more. Manuscripts submitted to the Journal with authors names included with have the authors' names removed by the Editor-in-Chief before being reviewed to ensure anonymity.

Typically, authors receive the reviewers' comments about a month after the submission of the manuscript. The Journal uses a constructive process for review and preparation of the manuscript, and encourages its reviewers to give supportive and extensive feedback on the requirements for improving the manuscript as well as guidance on how to make the amendments.

If the manuscript is deemed acceptable for publication, and reviewers' comments have been satisfactorily addressed, the manuscript is prepared for publication by the Copy Editor. The Copy Editor may correspond with the authors to check details, if required. Final publication is by discretion of the Editor-in-Chief. Final published form of the manuscript is via the Journal website (www.apjce.org), authors will be notified and sent a PDF copy of the final manuscript. There is no charge for publishing in APJCE and the Journal allows free open access for its readers.

Types of Manuscripts Sought by the Journal

Types of manuscripts the Journal accepts are primarily of two forms; *research reports* describing research into aspects of Cooperative Education and Work Integrated Learning/Education, and *topical discussion* articles that review relevant literature and give critical explorative discussion around a topical issue.

The Journal does also accept *best practice* papers but only if it present a unique or innovative practice of a Co-op/WIL program that is likely to be of interest to the broader Co-op/WIL community. The Journal also accepts a limited number of *Book Reviews* of relevant and recently published books.

Research reports should contain; an introduction that describes relevant literature and sets the context of the inquiry, a description and justification for the methodology employed, a description of the research findings-tabulated as appropriate, a discussion of the importance of the findings including their significance for practitioners, and a conclusion preferably incorporating suggestions for further research.

Topical discussion articles should contain a clear statement of the topic or issue under discussion, reference to relevant literature, critical discussion of the importance of the issues, and implications for other researchers and practitioners.

Asia-Pacific Journal of Cooperative Education
ISSN 1175-2882
www.apjce.org

EDITORIAL BOARD